

# Artificial Intelligence (AI)

Oguzhan (Ouz) Gencoglu

Tampere University of Technology  
Top Data Science

Turku Turbulence - 8 November 2016

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

# Tampere University of Technology



TAMPERE  
UNIVERSITY OF  
TECHNOLOGY

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

# Top Data Science



A Top Data Science Company

# Outline

## 1 What Exactly Is AI/Machine Learning?

- Some Terminology
- Quick History
- Recent Highlights

## 2 What Can Possibly Happen?

- Theories, Possibilities, Claims & Loads of Bullsh\*t
- Some Thoughts

## 3 But, How Do These Work?

- Teach Me Master
- There Is No Free Lunch!
- 2-minute Science
- Consequences

## 4 What Can Be Done?

## 5 Where Are We Heading?

# Outline

## 1 What Exactly Is AI/Machine Learning?

### ■ Some Terminology

- Quick History
- Recent Highlights

## 2 What Can Possibly Happen?

- Theories, Possibilities, Claims & Loads of Bullsh\*t
- Some Thoughts

## 3 But, How Do These Work?

- Teach Me Master
- There Is No Free Lunch!
- 2-minute Science
- Consequences

## 4 What Can Be Done?

## 5 Where Are We Heading?

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Some Terminology  
Quick History  
Recent Highlights

# Some Terminology

Data Science

Attempt to make sense out of data

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Some Terminology  
Quick History  
Recent Highlights

# Some Terminology

## Data Science

Attempt to make sense out of data

## Machine Learning (ML)

Subfield of *computer science*, algorithms that can learn without being explicitly programmed

# Some Terminology

## Data Science

Attempt to make sense out of data

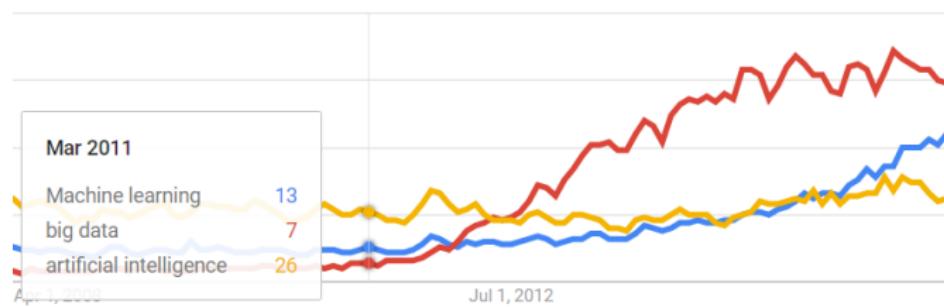
## Machine Learning (ML)

Subfield of *computer science*, algorithms that can learn without being explicitly programmed

## Artificial Intelligence (AI)

Machines mimicking *cognitive* functions of humans

# Some Terminology



# Outline

## 1 What Exactly Is AI/Machine Learning?

- Some Terminology
- Quick History
- Recent Highlights

## 2 What Can Possibly Happen?

- Theories, Possibilities, Claims & Loads of Bullsh\*t
- Some Thoughts

## 3 But, How Do These Work?

- Teach Me Master
- There Is No Free Lunch!
- 2-minute Science
- Consequences

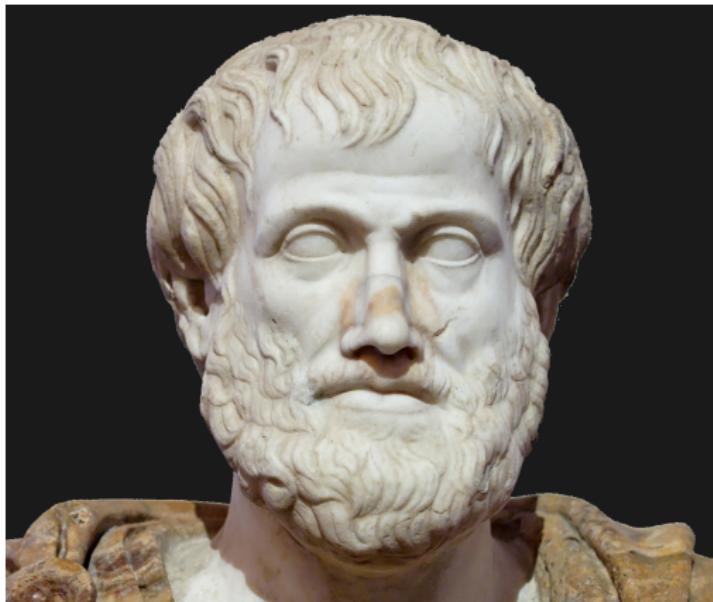
## 4 What Can Be Done?

## 5 Where Are We Heading?

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Some Terminology  
Quick History  
Recent Highlights

# Quick History



What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Some Terminology  
Quick History  
Recent Highlights

# Quick History



# Quick History

- Ancient Greek: Aristotle, first formal deductive reasoning system

# Quick History

- Ancient Greek: Aristotle, first formal deductive reasoning system
- 13th century: Al-Jazari, first programmable robot

# Quick History

- Ancient Greek: Aristotle, first formal deductive reasoning system
- 13th century: Al-Jazari, first programmable robot
- 17th century: Thomas Hobbes, mechanistic theory of thinking

# Quick History

- Ancient Greek: Aristotle, first formal deductive reasoning system
- 13th century: Al-Jazari, first programmable robot
- 17th century: Thomas Hobbes, mechanistic theory of thinking
- Descartes, Pascal, Leibniz ...

# Quick History

- Ancient Greek: Aristotle, first formal deductive reasoning system
- 13th century: Al-Jazari, first programmable robot
- 17th century: Thomas Hobbes, mechanistic theory of thinking
- Descartes, Pascal, Leibniz ...
- Charles Babbage (1832), the Analytical Engine

# Quick History

- Ancient Greek: Aristotle, first formal deductive reasoning system
- 13th century: Al-Jazari, first programmable robot
- 17th century: Thomas Hobbes, mechanistic theory of thinking
- Descartes, Pascal, Leibniz ...
- Charles Babbage (1832), the Analytical Engine
- George Boole (1854), binary algebra

# Quick History

- Ancient Greek: Aristotle, first formal deductive reasoning system
- 13th century: Al-Jazari, first programmable robot
- 17th century: Thomas Hobbes, mechanistic theory of thinking
- Descartes, Pascal, Leibniz ...
- Charles Babbage (1832), the Analytical Engine
- George Boole (1854), binary algebra
- 20th century - first half, Bertrand Russell, Alan Turing, Claude Shannon, Isaac Asimov

# Quick History

- Ancient Greek: Aristotle, first formal deductive reasoning system
- 13th century: Al-Jazari, first programmable robot
- 17th century: Thomas Hobbes, mechanistic theory of thinking
- Descartes, Pascal, Leibniz ...
- Charles Babbage (1832), the Analytical Engine
- George Boole (1854), binary algebra
- 20th century - first half, Bertrand Russell, Alan Turing, Claude Shannon, Isaac Asimov
- Modern programming languages, machine learning, robotics.

# Quick History

- Ancient Greek: Aristotle, first formal deductive reasoning system
- 13th century: Al-Jazari, first programmable robot
- 17th century: Thomas Hobbes, mechanistic theory of thinking
- Descartes, Pascal, Leibniz ...
- Charles Babbage (1832), the Analytical Engine
- George Boole (1854), binary algebra
- 20th century - first half, Bertrand Russell, Alan Turing, Claude Shannon, Isaac Asimov
- Modern programming languages, machine learning, robotics.
- ...

# Outline

## 1 What Exactly Is AI/Machine Learning?

- Some Terminology
- Quick History
- Recent Highlights

## 2 What Can Possibly Happen?

- Theories, Possibilities, Claims & Loads of Bullsh\*t
- Some Thoughts

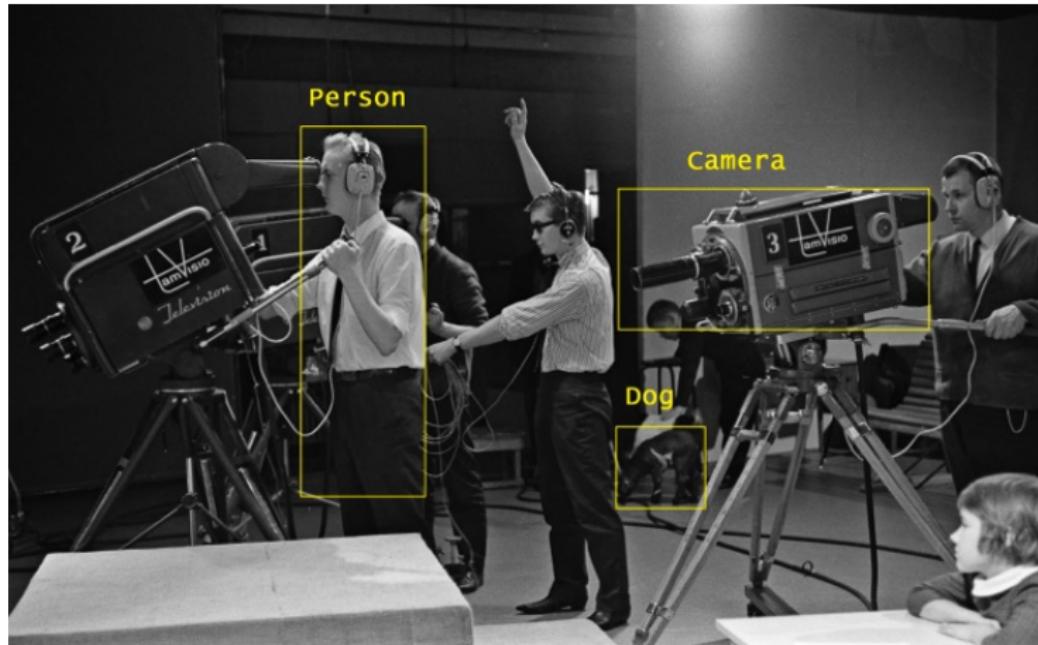
## 3 But, How Do These Work?

- Teach Me Master
- There Is No Free Lunch!
- 2-minute Science
- Consequences

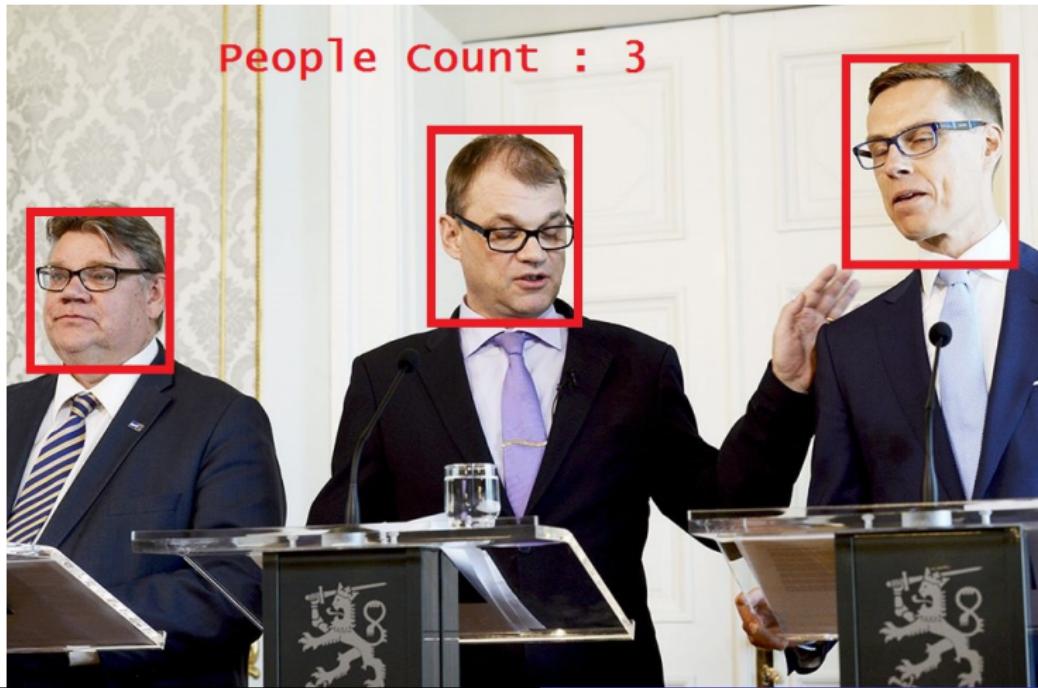
## 4 What Can Be Done?

## 5 Where Are We Heading?

# Generic Object Recognition



# People Count



# Age & Gender Estimation



# Person Identification



© PA

# Visual Description



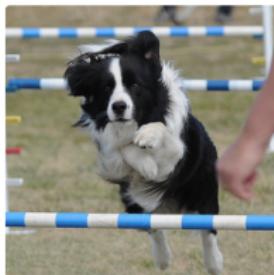
"man in black shirt is playing guitar."



"construction worker in orange safety vest is working on road."



"girl in pink dress is jumping in air"



"black and white dog jumps over bar"

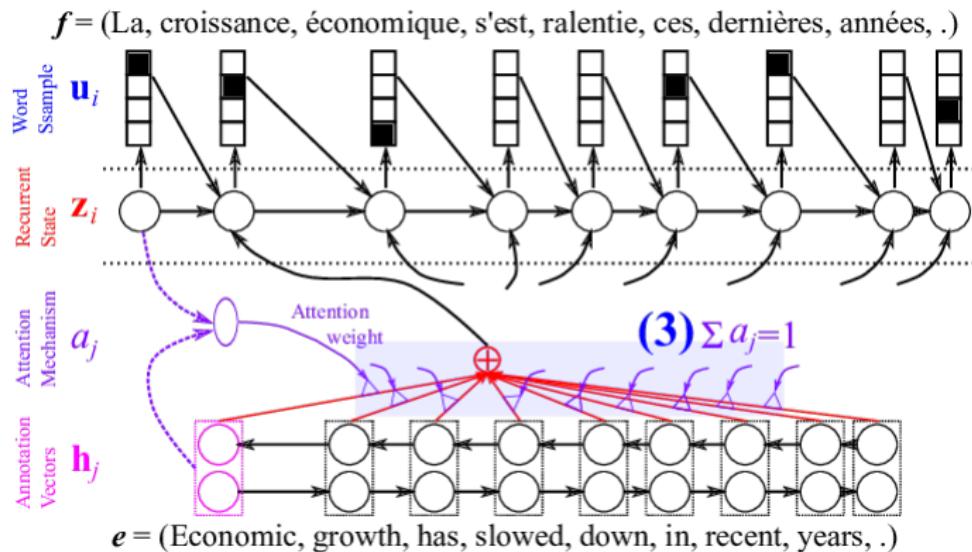
What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Some Terminology  
Quick History  
Recent Highlights

# Style Transfer



# Machine Translation



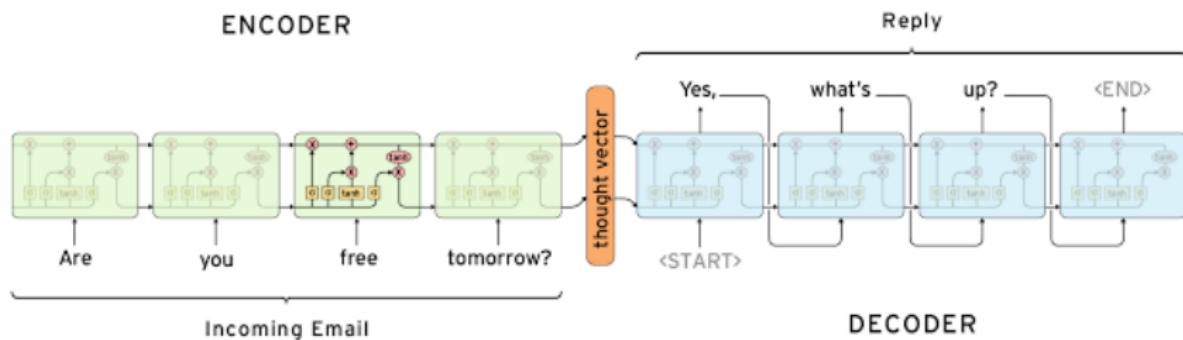
What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Some Terminology  
Quick History  
Recent Highlights

# AI in Games



# Chatbots



DeepDrumpf  
@DeepDrumpf



Follow

OK, it's amazing right now with ISIS, I tell you what? I don't want them to vote, the worst very social people. I love me.



What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Some Terminology  
Quick History  
Recent Highlights

# Image Super-resolution



What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Some Terminology  
Quick History  
Recent Highlights

# Image Generation



What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Some Terminology  
Quick History  
Recent Highlights

# Assessing Actions

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Some Terminology  
Quick History  
Recent Highlights

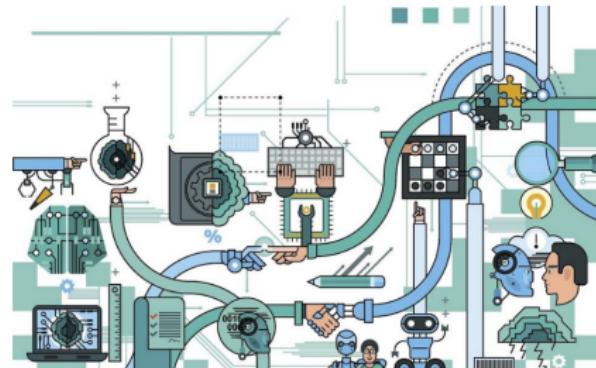
# Predicting Actions

# Topic Classification



TIETOJENKÄSITTELY | Tapio Ikkala | 23.8. klo 11:34

## Koneoppimiskilpailun finalistit Tampereelta - Yhden miehen tiimi haastaa IBM:n Watsonin



What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Some Terminology  
Quick History  
Recent Highlights

# Clinical Decision Support



What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Some Terminology  
Quick History  
Recent Highlights

# Self-driving Vehicles



What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Some Terminology  
Quick History  
Recent Highlights

# Other Examples

- Speech Recognition

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Some Terminology  
Quick History  
Recent Highlights

# Other Examples

- Speech Recognition
- Automated Workforce

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Some Terminology  
Quick History  
Recent Highlights

# Other Examples

- Speech Recognition
- Automated Workforce
- Predicting Epidemics From Social Media

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Some Terminology  
Quick History  
Recent Highlights

# Other Examples

- Speech Recognition
- Automated Workforce
- Predicting Epidemics From Social Media
- Targeted Advertising

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Some Terminology  
Quick History  
Recent Highlights

# Other Examples

- Speech Recognition
- Automated Workforce
- Predicting Epidemics From Social Media
- Targeted Advertising
- Anonymization & Leveraging Security

# Other Examples

- Speech Recognition
- Automated Workforce
- Predicting Epidemics From Social Media
- Targeted Advertising
- Anonymization & Leveraging Security
- De-anonymization, Fraud Detection, Criminal Identification

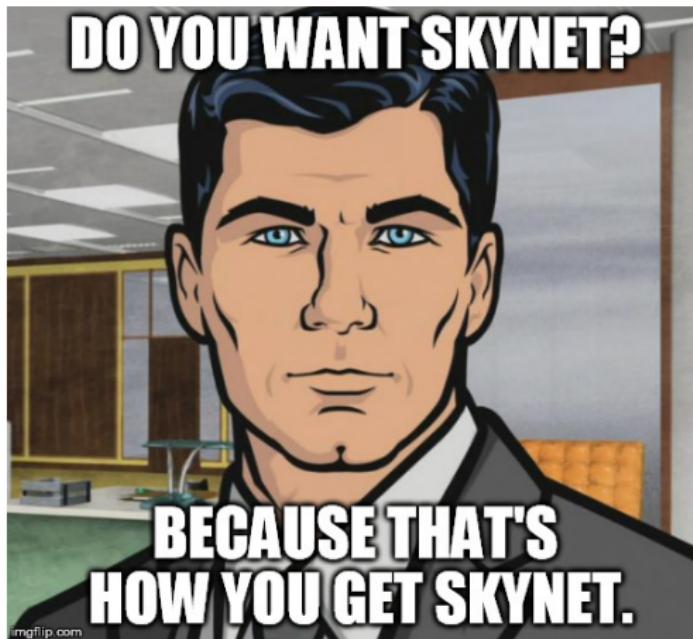
# Other Examples

- Speech Recognition
- Automated Workforce
- Predicting Epidemics From Social Media
- Targeted Advertising
- Anonymization & Leveraging Security
- De-anonymization, Fraud Detection, Criminal Identification
- ...

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Theories, Possibilities, Claims & Loads of Bullsh\*t  
Some Thoughts

# What Can Possibly Happen?



# Outline

## 1 What Exactly Is AI/Machine Learning?

- Some Terminology
- Quick History
- Recent Highlights

## 2 What Can Possibly Happen?

- Theories, Possibilities, Claims & Loads of Bullsh\*t
- Some Thoughts

## 3 But, How Do These Work?

- Teach Me Master
- There Is No Free Lunch!
- 2-minute Science
- Consequences

## 4 What Can Be Done?

## 5 Where Are We Heading?

# Theories, Possibilities, Claims & Loads of Bullsh\*t

## Claim

- Superintelligence by 2100 is inevitable!
- Superintelligence by 2100 is impossible!

# Theories, Possibilities, Claims & Loads of Bullsh\*t

## Claim

- AI will turn evil!
- AI will turn conscious!

# Theories, Possibilities, Claims & Loads of Bullsh\*t

## Claim

- AI will turn evil!
- AI will turn conscious!

## Actual Worry

- AI becoming very competent with its own goals.

# Theories, Possibilities, Claims & Loads of Bullsh\*t

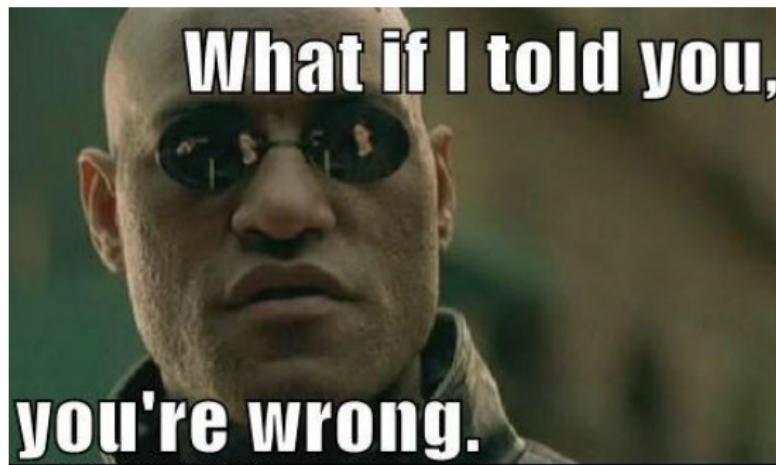
## Claim

- A creature can not acquire higher intelligence than its creator.

# Theories, Possibilities, Claims & Loads of Bullsh\*t

## Claim

- A creature can not acquire higher intelligence than its creator.



# Theories, Possibilities, Claims & Loads of Bullsh\*t

## Claim

- Robots will replace labour. Low-paid work or unemployment seem to be the only possible options.

# Theories, Possibilities, Claims & Loads of Bullsh\*t

## Claim

- Robots will replace labour. Low-paid work or unemployment seem to be the only possible options.



# Outline

## 1 What Exactly Is AI/Machine Learning?

- Some Terminology
- Quick History
- Recent Highlights

## 2 What Can Possibly Happen?

- Theories, Possibilities, Claims & Loads of Bullsh\*t
- Some Thoughts

## 3 But, How Do These Work?

- Teach Me Master
- There Is No Free Lunch!
- 2-minute Science
- Consequences

## 4 What Can Be Done?

## 5 Where Are We Heading?

**Everyone can enjoy a life of luxurious leisure if the machine-produced wealth is shared, or most people can end up miserably poor if the machine-owners successfully lobby against wealth redistribution. So far, the trend seems to be toward the second option, with technology driving ever-increasing inequality.**

*Stephen Hawking*

**An extremely intelligent future AI will probably develop a drive to survive and acquire more resources as a step toward accomplishing whatever goal it has, because surviving and having more resources will increase its chances of accomplishing that other goal.**

*Steve Omohundro*

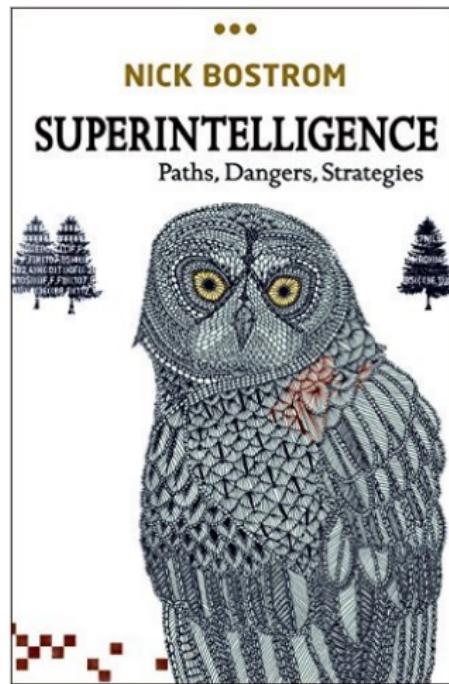
What Exactly Is AI/Machine Learning?  
**What Can Possibly Happen?**  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Theories, Possibilities, Claims & Loads of Bullsh\*t  
**Some Thoughts**

**What level of sophistication will artificial intelligences need to attain before we consider that they deserve legal rights?**

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Theories, Possibilities, Claims & Loads of Bullsh\*t  
Some Thoughts



What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
**But, How Do These Work?**  
What Can Be Done?  
Where Are We Heading?  
Conclusion

**Teach Me Master**  
There Is No Free Lunch!  
2-minute Science  
Consequences

# Outline

## 1 What Exactly Is AI/Machine Learning?

- Some Terminology
- Quick History
- Recent Highlights

## 2 What Can Possibly Happen?

- Theories, Possibilities, Claims & Loads of Bullsh\*t
- Some Thoughts

## 3 But, How Do These Work?

- Teach Me Master
- There Is No Free Lunch!
- 2-minute Science
- Consequences

## 4 What Can Be Done?

## 5 Where Are We Heading?

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
**But, How Do These Work?**  
    What Can Be Done?  
    Where Are We Heading?  
    Conclusion

Teach Me Master  
There Is No Free Lunch!  
2-minute Science  
Consequences

## But, How Do These Work?

(Most) Machine learning algorithms learn from examples, i.e.,  
supervised learning.

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
**But, How Do These Work?**  
    What Can Be Done?  
    Where Are We Heading?  
    Conclusion

Teach Me Master  
There Is No Free Lunch!  
2-minute Science  
Consequences

## But, How Do These Work?

(Most) Machine learning algorithms learn from examples, i.e.,  
supervised learning.

Accuracies tend to increase with more data.

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
**But, How Do These Work?**  
    What Can Be Done?  
    Where Are We Heading?  
    Conclusion

Teach Me Master  
There Is No Free Lunch!  
2-minute Science  
Consequences

## But, How Do These Work?

(Most) Machine learning algorithms learn from examples, i.e.,  
supervised learning.

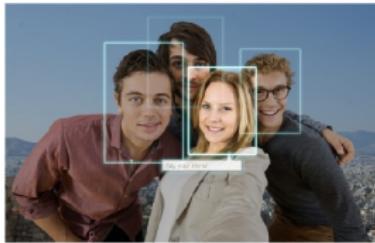
Accuracies tend to increase with more data.

**YOU ARE THE TEACHER!**

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
**But, How Do These Work?**  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Teach Me Master  
There Is No Free Lunch!  
2-minute Science  
Consequences

# But, How Do These Work?



## Other Movies You Might Enjoy

Balls of Fury has been added to your DVI at position 454.

This movie is available.

[Move To Top Of My Queue](#)

[Return To Michael Ian Black Movies](#)

**Balls of Fury**  
Demolition Man

**Add**

**Add**

5 stars

Mr. Woodcock

Deuce Bigalow: European Gigolo

Delta Force

Date Movie

**Add**

**Add**

**Add**

**Add**

# Outline

## 1 What Exactly Is AI/Machine Learning?

- Some Terminology
- Quick History
- Recent Highlights

## 2 What Can Possibly Happen?

- Theories, Possibilities, Claims & Loads of Bullsh\*t
- Some Thoughts

## 3 But, How Do These Work?

- Teach Me Master
- **There Is No Free Lunch!**
- 2-minute Science
- Consequences

## 4 What Can Be Done?

## 5 Where Are We Heading?

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
**But, How Do These Work?**  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Teach Me Master  
There Is No Free Lunch!  
2-minute Science  
Consequences

# No Free Lunch Theorem

There is always a trade-off

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
**But, How Do These Work?**  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Teach Me Master  
There Is No Free Lunch!  
2-minute Science  
Consequences

# No Free Lunch Theorem

There is always a trade-off

- Complexity **vs.** Interpretability & Transparency

# No Free Lunch Theorem

There is always a trade-off

- Complexity **vs.** Interpretability & Transparency
- Accuracy **vs.** Computation Power

# No Free Lunch Theorem

There is always a trade-off

- Complexity **vs.** Interpretability & Transparency
- Accuracy **vs.** Computation Power
- Specificity **vs.** Sensitivity

# No Free Lunch Theorem

There is always a trade-off

- Complexity **vs.** Interpretability & Transparency
- Accuracy **vs.** Computation Power
- Specificity **vs.** Sensitivity
- Bias **vs.** Variance

# No Free Lunch Theorem

There is always a trade-off

- Complexity **vs.** Interpretability & Transparency
- Accuracy **vs.** Computation Power
- Specificity **vs.** Sensitivity
- Bias **vs.** Variance
- ...

# Outline

## 1 What Exactly Is AI/Machine Learning?

- Some Terminology
- Quick History
- Recent Highlights

## 2 What Can Possibly Happen?

- Theories, Possibilities, Claims & Loads of Bullsh\*t
- Some Thoughts

## 3 But, How Do These Work?

- Teach Me Master
- There Is No Free Lunch!
- **2-minute Science**
- Consequences

## 4 What Can Be Done?

## 5 Where Are We Heading?

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
**But, How Do These Work?**  
    What Can Be Done?  
    Where Are We Heading?  
    Conclusion

Teach Me Master  
There Is No Free Lunch!  
2-minute Science  
Consequences

# An Example

Certain patterns occur frequently in nature.

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
**But, How Do These Work?**  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Teach Me Master  
There Is No Free Lunch!  
2-minute Science  
Consequences

# An Example

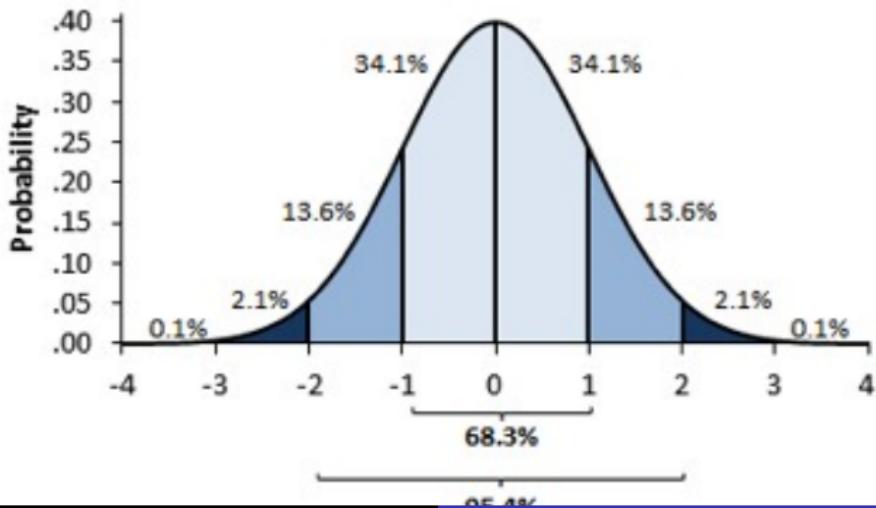
Certain patterns occur frequently in nature.

Normal (Gaussian) Distribution

# An Example

Certain patterns occur frequently in nature.

## Normal (Gaussian) Distribution



# Normal Distribution

## ■ Velocity of Molecules in a Gas

---

<sup>1</sup>Huxley, Julian S. (1932). Problems of Relative Growth.

<sup>2</sup>Oosterbaan, Roland J. (1994). "Chapter 6: Frequency and Regression Analysis of Hydrologic Data". In Ritzema, Henk P. Drainage Principles and Applications

<sup>3</sup>Andersen, Torben G., et al. (2001) "The distribution of realized stock return volatility." Journal of financial economics 61.1

# Normal Distribution

- Velocity of Molecules in a Gas
- Human Height<sup>1</sup>

---

<sup>1</sup>Huxley, Julian S. (1932). Problems of Relative Growth.

<sup>2</sup>Oosterbaan, Roland J. (1994). "Chapter 6: Frequency and Regression Analysis of Hydrologic Data". In Ritzema, Henk P. Drainage Principles and Applications

<sup>3</sup>Andersen, Torben G., et al. (2001) "The distribution of realized stock return volatility." Journal of financial economics 61.1

# Normal Distribution

- Velocity of Molecules in a Gas
- Human Height<sup>1</sup>
- Monthly Rainfall<sup>2</sup>

---

<sup>1</sup>Huxley, Julian S. (1932). Problems of Relative Growth.

<sup>2</sup>Oosterbaan, Roland J. (1994). "Chapter 6: Frequency and Regression Analysis of Hydrologic Data". In Ritzema, Henk P. Drainage Principles and Applications

<sup>3</sup>Andersen, Torben G., et al. (2001) "The distribution of realized stock return volatility." Journal of financial economics 61.1

# Normal Distribution

- Velocity of Molecules in a Gas
- Human Height<sup>1</sup>
- Monthly Rainfall<sup>2</sup>
- Stock Return Volatility<sup>3</sup>

---

<sup>1</sup>Huxley, Julian S. (1932). Problems of Relative Growth.

<sup>2</sup>Oosterbaan, Roland J. (1994). "Chapter 6: Frequency and Regression Analysis of Hydrologic Data". In Ritzema, Henk P. Drainage Principles and Applications

<sup>3</sup>Andersen, Torben G., et al. (2001) "The distribution of realized stock return volatility." Journal of financial economics 61.1

# Normal Distribution

- Velocity of Molecules in a Gas
- Human Height<sup>1</sup>
- Monthly Rainfall<sup>2</sup>
- Stock Return Volatility<sup>3</sup>

Why?  $\implies$  Central Limit Theorem

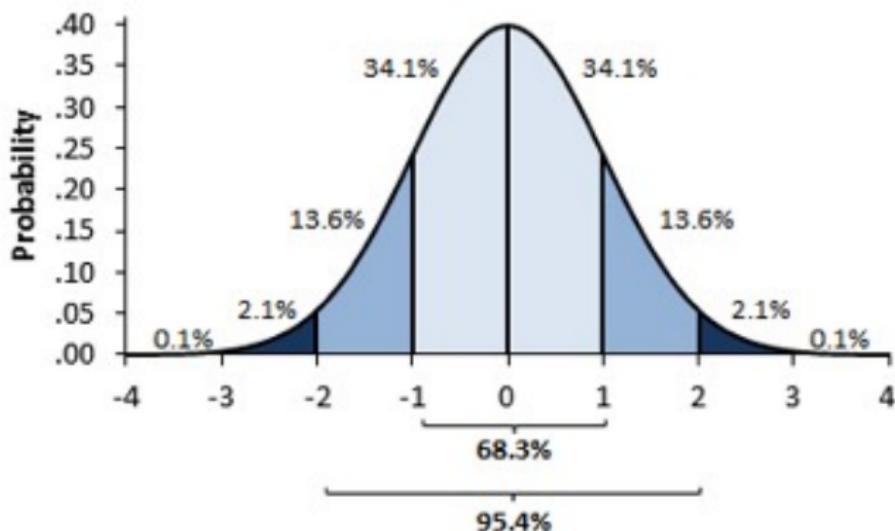
---

<sup>1</sup>Huxley, Julian S. (1932). Problems of Relative Growth.

<sup>2</sup>Oosterbaan, Roland J. (1994). "Chapter 6: Frequency and Regression Analysis of Hydrologic Data". In Ritzema, Henk P. Drainage Principles and Applications

<sup>3</sup>Andersen, Torben G., et al. (2001) "The distribution of realized stock return volatility." Journal of financial economics 61.1

## Inevitable Nature of Things



There will be outliers, anomalies and under-represented groups/subsets!

# Outline

## 1 What Exactly Is AI/Machine Learning?

- Some Terminology
- Quick History
- Recent Highlights

## 2 What Can Possibly Happen?

- Theories, Possibilities, Claims & Loads of Bullsh\*t
- Some Thoughts

## 3 But, How Do These Work?

- Teach Me Master
- There Is No Free Lunch!
- 2-minute Science
- **Consequences**

## 4 What Can Be Done?

## 5 Where Are We Heading?

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
**But, How Do These Work?**  
    What Can Be Done?  
    Where Are We Heading?  
    Conclusion

Teach Me Master  
There Is No Free Lunch!  
2-minute Science  
**Consequences**

## Consequences

# Supervised Algorithms

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
**But, How Do These Work?**  
    What Can Be Done?  
    Where Are We Heading?  
    Conclusion

Teach Me Master  
There Is No Free Lunch!  
2-minute Science  
**Consequences**

## Consequences

# Supervised Algorithms + Biased Data

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
**But, How Do These Work?**  
    What Can Be Done?  
    Where Are We Heading?  
    Conclusion

Teach Me Master  
There Is No Free Lunch!  
2-minute Science  
**Consequences**

## Consequences

# Supervised Algorithms

+

# Biased Data

=

# Biased Algorithms

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
**But, How Do These Work?**  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Teach Me Master  
There Is No Free Lunch!  
2-minute Science  
Consequences

# Ugly Results

If the training data reflect existing social biases against a minority, the algorithm is likely to incorporate these biases.

---

<sup>4</sup><http://www.independent.co.uk/life-style/gadgets-and-tech/news/googles-algorithm-shows-prestigious-job-ads-to-men-but-not-to-women-10372166.html>

<sup>5</sup>[http://www.huffingtonpost.com/2015/04/10/google-image-gender-bias\\_n\\_7036414.html](http://www.huffingtonpost.com/2015/04/10/google-image-gender-bias_n_7036414.html) 

## Ugly Results

If the training data reflect existing social biases against a minority, the algorithm is likely to incorporate these biases.

Google's algorithms showing prestigious job ads to men, but not to women.<sup>4</sup>

---

<sup>4</sup><http://www.independent.co.uk/life-style/gadgets-and-tech/news/googles-algorithm-shows-prestigious-job-ads-to-men-but-not-to-women-10372166.html>

<sup>5</sup>[http://www.huffingtonpost.com/2015/04/10/google-image-gender-bias\\_n\\_7036414.html](http://www.huffingtonpost.com/2015/04/10/google-image-gender-bias_n_7036414.html)

## Ugly Results

If the training data reflect existing social biases against a minority, the algorithm is likely to incorporate these biases.

Google's algorithms showing prestigious job ads to men, but not to women.<sup>4</sup>

Google image search in USA for “CEO” produced 11 % women, even though 27 % of United States chief executives are women.<sup>5</sup>

---

<sup>4</sup><http://www.independent.co.uk/life-style/gadgets-and-tech/news/googles-algorithm-shows-prestigious-job-ads-to-men-but-not-to-women-10372166.html>

<sup>5</sup>[http://www.huffingtonpost.com/2015/04/10/google-image-gender-bias\\_n\\_7036414.html](http://www.huffingtonpost.com/2015/04/10/google-image-gender-bias_n_7036414.html)

# Ugly Results

Consider a loan application from a bank. Numerous attributes can be an input to a ML algorithm:

- Age
- Field of Work
- Income
- Gender
- Marital Status
- Residence District
- Number of Children
- Transaction History
- Race
- ...

# Ugly Results

Consider a loan application from a bank. Numerous attributes can be an input to a ML algorithm:

- Age
- Field of Work
- Income
- **Gender !**
- Marital Status
- Residence District
- Number of Children
- Transaction History
- **Race !**
- ...

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
**But, How Do These Work?**  
What Can Be Done?  
Where Are We Heading?  
Conclusion

Teach Me Master  
There Is No Free Lunch!  
2-minute Science  
Consequences

# Ugly Results

Anything else?

# Ugly Results

Anything else?

- Age
- Field of Work
- Income
- **Gender !**
- Marital Status
- **Residence District !**
- Number of Children
- Transaction History
- **Race !**
- ...

Biases can be transferred through correlated attributes!



# Discrimination and Privacy-aware ML

Bias-free machine learning research should be encouraged.

# Transparent Design

Is the decision explainable enough?

# Transparent Design

Is the decision explainable enough?

Which attributes (features) have been used?

# Transparent Design

Is the decision explainable enough?

Which attributes (features) have been used?

Importance of attributes inferred by the algorithm?

# Transparent Reporting

New ML algorithm predicts whether one has lung cancer or not with 90% accuracy.

# Transparent Reporting

New ML algorithm predicts whether one has lung cancer or not with 90% accuracy.

**Overall accuracy is not enough.**

# Transparent Reporting

New ML algorithm predicts whether one has lung cancer or not with 90% accuracy.

**Overall accuracy is not enough.**

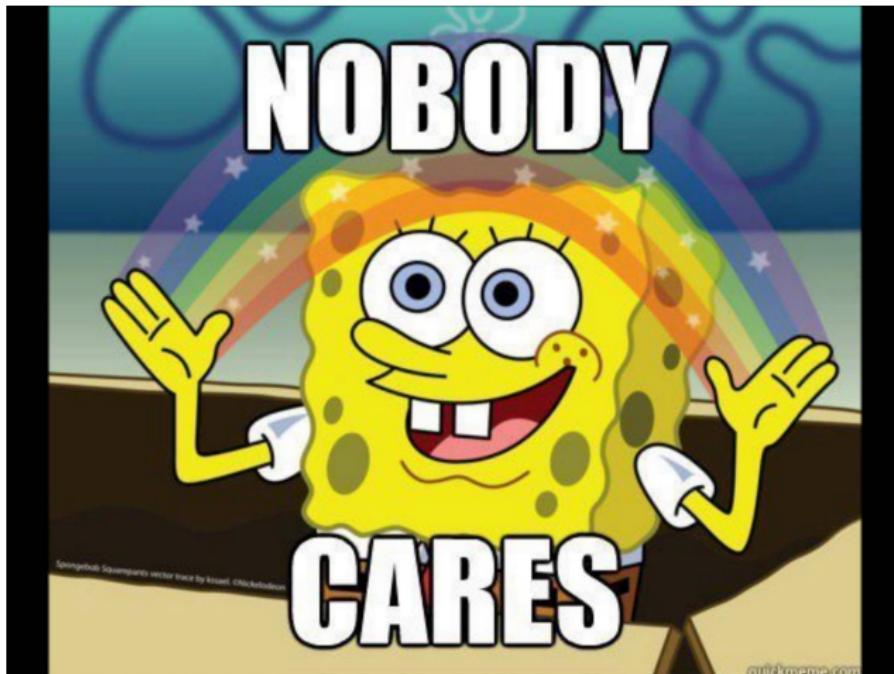
**Accuracies on different subgroups, false positive & true negative rates.**

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

# Current State of Awareness

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

## Current State of Awareness



## Box Is Getting Blacker

Black-box models (e.g. neural networks) are extremely popular!

## Box Is Getting Blacker

Black-box models (e.g. neural networks) are extremely popular!

*“If it works, I don’t care why!”* approach.

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

# Conclusion

AI is coming (already here actually):

Deal with it!

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

# Conclusion

AI is coming (already here actually):

Deal with it!

Things will change:

Let's research on creating beneficial intelligence!

What Exactly Is AI/Machine Learning?  
What Can Possibly Happen?  
But, How Do These Work?  
What Can Be Done?  
Where Are We Heading?  
Conclusion

